

DYNAMIC SYSTEM CALIBRATION BY SYSTEM IDENTIFICATION METHODS

R. Ergon*, D. Di Ruscio*

*Telemark Institute of Technology
N-3914 Porsgrunn, Norway

Fax: + 47 35 57 52 50 and e-mail: Rolf.Ergon@hit.no

Keywords: Estimation, system identification, multivariate calibration

Abstract

Primary output variables from industrial processes can be estimated from known input variables and secondary process measurements. As a basis for this, the dynamic predictor has to be identified from data collected during a calibration experiment. In this paper, the theoretical basis for this is investigated, and a systematic experimental method is proposed.

1 Introduction

In many industrial process plants, vital output variables are not available on-line. In such cases, an indirect measure may be used to infer the process variables of interest, see e.g. [1]. Typically, product qualities y_1 are inferred from process measurements y_2 . In some cases, several collinear secondary measurements and multivariable calibration methods are used to estimate primary variables [2]. The success of these inferential methods depends on a good knowledge of the static relation between the measured and estimated variables. That relation can be established either through first principle modeling or by use of data from a calibration/identification experiment, or possibly by a combination of both these methods.

In this paper we study the use of inferential methods covering also dynamical relations, focusing on the empirical approach and use of system identification methods. The purpose, then, is to infer primary process variables y_1 from known inputs u (manipulated or measured) and secondary process measurements y_2 . That is, given input-output measurements from controlled experiments where also y_1 is available, we want to establish the dynamical model that is necessary for estimation of vital output variables that are not available on-line during normal operation. In such an inferential method, also y_2 is used as an input signal, and could thus be included in u . For theoretical analysis and clarity of presentation, though, we will separate between the input vector u and the measurement vector y_2 .

The specific problem dealt with in the paper, is the use of both independent variables u and dependant variables y_2 as inputs in a system identification procedure. From a system identification point of view, that is a very natural idea, and certainly not new [3]. The basic idea is that for output estimation purposes, knowledge of the system model as such is not necessary. What is needed are the dynamical relations between the known input signals u , the available measurements y_2 and the output variables y_1 , and these relations can often be identified with better accuracy than the relations between u and y_1 alone. The reason for this is that process disturbances and noise entering early in the process, will be indirectly measured by available process measurements later in the process.

The aim of the present work is to investigate the effects of utilizing the available y_2 measurements as input signals in order to estimate the primary variables y_1 , and this is carried out by use of a Kalman filtering approach. In order to find the appropriate Kalman gains, it is then necessary to specify an output error model. This also means that the feedback path from y_1 to y_2 is broken, resulting in an open loop identification problem. A systematic experimental method for optimal utilization of the y_2 measurements is also proposed. The theoretical analysis and the proposed experimental method are limited to stable systems. Thus, for unstable processes, stabilizing controllers have to be included in the system model.

The proposed method can be seen as a calibration procedure, that relates the available process inputs u and measurements y_2 to the output variables y_1 . Since the system generally is dynamic in nature, it seems appropriate to use the term "dynamic system calibration". It should be emphasized though, that this calibration is based on well established system identification theory, see e.g. [4].

The paper is organized in the following way: In section 2 the theoretical basis is established, a bias analysis is performed and deterministic and perfect measurement cases are discussed. In section 3 an experimental calibration method is proposed, while section 4 gives some simulation results. Concluding remarks are given in section 5.

2 Theory

2.1 Theoretical model

Consider the discrete system model

$$x_{k+1} = Ax_k + Bu_k + Gv_k \quad (1a)$$

$$y_{1,k} = C_1x_k + D_1u_k + w_{1,k} \quad (1b)$$

$$y_{2,k} = C_2x_k + D_2u_k + w_{2,k}, \quad (1c)$$

where v , w_1 and w_2 are white and independent process and measurement noise vectors. The reason why w_1 and w_2 must be independent is given below.

The model (1) can be expressed in the ordinary innovation form [4], given by the following equations, where $K = \begin{bmatrix} K_1 & K_2 \end{bmatrix}$ is the Kalman gain:

$$\bar{x}_{k+1} = A\bar{x}_k + Bu_k + \begin{bmatrix} K_1 & K_2 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}_k \quad (2a)$$

$$y_{1,k} = C_1\bar{x}_k + D_1u_k + e_{1,k} \quad (2b)$$

$$y_{2,k} = C_2\bar{x}_k + D_2u_k + e_{2,k}. \quad (2c)$$

The classical optimal predictor with all measurements available will then be

$$\begin{aligned} \bar{x}_{k+1} &= (A - K_1C_1 - K_2C_2)\bar{x}_k \\ &\quad + (B - K_1D_1 - K_2D_2)u_k \\ &\quad + K_1y_{1,k} + K_2y_{2,k} \end{aligned} \quad (3a)$$

$$\bar{y}_{1,k} = C_1\bar{x}_k + D_1u_k \quad (3b)$$

$$\bar{y}_{2,k} = C_2\bar{x}_k + D_2u_k. \quad (3c)$$

When only the y_2 measurements are available, the optimal predictor for y_1 is

$$\begin{aligned} \bar{x}_{k+1} &= (A - K_2^{OE}C_2)\bar{x}_k \\ &\quad + (B - K_2^{OE}D_2)u_k + K_2^{OE}y_{2,k} \end{aligned} \quad (4a)$$

$$\bar{y}_{1,k} = C_1\bar{x}_k + D_1u_k. \quad (4b)$$

Note that we here assume $K_1 = 0$ by definition, which with y_2 used as input signal results in an output error (OE) model. When this predictor is to be based on system identification, we have two choices. One is first to identify (2) with $K_1 = 0$, and then construct the predictor (4). For complex systems with a number of secondary measurements y_2 , that is a difficult task [3]. The other and more appealing choice, especially with only one or a few primary measurements y_1 , is to set $K_1 = 0$ and reorganize (2) in the following way before identification:

$$\begin{aligned} \bar{x}_{k+1} &= (A - K_2^{OE}C_2)\bar{x}_k \\ &\quad + (B - K_2^{OE}D_2)u_k + K_2^{OE}y_{2,k} \end{aligned} \quad (5a)$$

$$y_{1,k} = C_1\bar{x}_k + D_1u_k + e_{1,k}. \quad (5b)$$

This partitioned output error innovation form then gives the estimation relation that has to be identified through the calibration experiment. Note, however, that we now

must require that e_1 is independent of the input y_2 , and this is why we earlier specified that w_1 and w_2 must be independent.

For theoretical considerations, we can determine K_2^{OE} by first solving the algebraic Riccati equation

$$P = APA^T + GR_vG^T - APC_2^T(C_2PC_2^T + R_{w_2})^{-1}C_2PA^T, \quad (6)$$

where $P = E\{(x_k - \bar{x}_k)(x_k - \bar{x}_k)^T\}$ while R_v and R_{w_2} are the process and measurement noise covariance matrices [5]. We then find the Kalman gain as

$$K_2^{OE} = APC_2^T(C_2PC_2^T + R_{w_2})^{-1}. \quad (7)$$

Example 1 Consider the following pure delay system:

$$x_{1,k+1} = x_{2,k} \quad (8)$$

$$x_{2,k+1} = u_k + v_k$$

$$y_{1,k} = x_{1,k} + w_{1,k}$$

$$y_{2,k} = x_{2,k} + w_{2,k}.$$

The theoretical predictor can be determined by first finding K_2^{OE} from (6) and (7). Due to the pure delay structure, the predictor (4) can then be reorganized into the input-output form

$$\bar{y}_{1,k} = \frac{1/r_v}{1/r_v + 1/r_2}u_{k-2} + \frac{1/r_2}{1/r_v + 1/r_2}y_{2,k-1} \quad (9)$$

where r_v and r_2 are the variances of v and w_2 .

This shows that the information in u and y_2 is utilized in an optimal way, considering the noise levels. ■

Example 2 Consider the 2. order system

$$\begin{aligned} x_{1,k+1} &= a_{11}x_{1,k} + a_{12}x_{2,k} + b_1u_k \\ x_{2,k+1} &= a_{21}x_{1,k} + a_{22}x_{2,k} + b_2u_k + v_k \end{aligned} \quad (10)$$

$$y_{1,k} = x_{1,k} + w_{1,k}$$

$$y_{2,k} = x_{2,k} + w_{2,k}.$$

Two simple cases can be easily calculated:

a) No process noise, i.e. $r_v/r_2 \rightarrow 0$, will result in $K_2^{OE} = 0$. With u as input and y_1 as output, we will then be able to identify the 2. order predictor

$$\begin{aligned} \bar{x}_{1,k+1} &= a_{11}\bar{x}_{1,k} + a_{12}\bar{x}_{2,k} + b_1u_k \\ \bar{x}_{2,k+1} &= a_{21}\bar{x}_{1,k} + a_{22}\bar{x}_{2,k} + b_2u_k \\ \bar{y}_{1,k} &= \bar{x}_{1,k}. \end{aligned} \quad (11)$$

b) With perfect measurements y_2 , i.e. $r_v/r_2 \rightarrow \infty$, the optimal predictor \bar{y}_1 will be found by ignoring the information in \bar{x}_2 , and rely on y_2 and the direct coupling from u to \bar{x}_1 (see Theorem 1 in section 2.4). We will therefore get $K_2^{OE} = \begin{bmatrix} a_{12} & k_{22}^{OE} \end{bmatrix}^T$ and from (5) the 2. order system

$$\begin{aligned} \bar{x}_{1,k+1} &= a_{11}\bar{x}_{1,k} + b_1u_k + a_{12}y_{2,k} \\ \bar{x}_{2,k+1} &= a_{21}\bar{x}_{1,k} + (a_{22} - k_{22}^{OE})\bar{x}_{2,k} \\ &\quad + b_2u_k + k_{22}^{OE}y_{2,k} \\ y_{1,k} &= \bar{x}_{1,k} + e_{1,k}. \end{aligned} \quad (12)$$

From this we see that the influence on y_1 from u and y_2 through \bar{x}_2 is decoupled, and we can therefore only identify the 1. order predictor

$$\begin{aligned}\bar{x}_{1,k+1} &= a_{11}\bar{x}_{1,k} + b_1u_k + a_{12}y_{2,k} \\ \bar{y}_{1,k} &= \bar{x}_{1,k}.\end{aligned}\quad (13)$$

Other values of r_v/r_2 will give intermediate results. ■

2.2 Bias analysis

The system (2) could be identified by use of a prediction error identification method. We would then employ the predictor (3), and assuming that the model structure is rich enough to cover the true system the result would be asymptotically (when the number of samples $N \rightarrow \infty$) unbiased parameter estimates, including the parameters in K_1 and K_2 as theoretically given by the Kalman filter formalism [6]. Theoretically, we would then find the optimal predictor (4) with $K_1 \rightarrow 0$ when $r_1 = Ew_1^2 \rightarrow \infty$ (or $r_i \rightarrow \infty$, $i = 1, 2, \dots, m$ with a m dimensional y_1 signal), and the parameter estimates would still be asymptotically unbiased. The practical way of doing this is to specify $K_1 = 0$, and utilize (4) instead of (3) during the identification.

When we are identifying (5), we will employ exactly the same predictor (4) as when identifying (2) with $K_1 = 0$. The only differences are that we use a simpler criterion function $V_1 = f_1(y_1 - \bar{y}_1)$ instead of $V_2 = f_2(y_1 - \bar{y}_1, y_2 - \bar{y}_2)$, and that we treat $A - K_2^{OE}C_2$ and $B - K_2^{OE}D_2$ as single matrices. Identification of (5) by use of a prediction error method will therefore result in asymptotically unbiased parameter estimates, including K_2^{OE} as given by (6) and (7). A more formal proof of this is given in [15].

2.3 Feedback and identifiability analysis

Using transfer functions $G(q^{-1})$ and $H(q^{-1})$, where q^{-1} is the unit delay operator, the model (2) can be expressed as

$$y_1 = G_1u + H_{11}e_1 + H_{12}e_2 \quad (14a)$$

$$y_2 = G_2u + H_{21}e_1 + H_{22}e_2. \quad (14b)$$

This gives

$$y_2 = G_2u + H_{21}H_{11}^{-1}(y_1 - G_1u - H_{12}e_2) + H_{22}e_2, \quad (15)$$

which means that we will generally have feedback from y_1 to y_2 , with some possible identifiability problems as a consequence [6]. In the present case, however, we identify (5) with $K_1 = 0$, and we will then also have $H_{21} = 0$. The feedback is therefore broken, and the system is then basically identifiable as long as u is persistently exciting of appropriate order. We must, however, also require that the Kalman gain K_2^{OE} theoretically can be determined by (6) and (7), and that the underlying Kalman filter is stable. This is always the case when we have some process noise, i.e. when $R_v > 0$, and at the same time $(A, G\sqrt{R_v})$ is stabilizable and (A, C_2) is detectable [7].

2.4 The deterministic case

It is well known [7], that the Kalman gain K in (2) cannot be determined in a pure deterministic case, that is when both the process noise v and the measurement noise w_1 and w_2 are zero. In the present case, this means that the Kalman gain K_2^{OE} in (5) cannot be identified.

In such a case, there is no need to utilize the information in the y_2 measurements, and the natural solution is then to identify a model with u as input and y_1 as output signals. Without noise, this can be solved as an ordinary linear regression problem using a least sum of squares method, that is, by identifying an ARX model [8].

Overlooking that we have a noise free case at hand, we would in the present context attempt to identify a model with both u and y_2 as input signals, resulting in a rank deficient data matrix. There are two general solutions to this problem:

- a) Identify a model with only u as input signal.
- b) Identify an appropriately reduced model, possibly with only y_2 as input signal. An experimental method for finding such a model is proposed in section 3.

A general treatment of the deterministic case is given in [15].

2.5 The asymptotic perfect measurement case

In the ideal situation, we have noise free y_2 measurements. In order to analyze such an asymptotic perfect measurement case, we use the model (1), partitioned in the following way:

$$\begin{aligned}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_{k+1} &= \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_k \\ &+ \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix} u_k + \begin{bmatrix} G_1 \\ G_2 \\ G_3 \end{bmatrix} v_k\end{aligned}\quad (16a)$$

$$\begin{aligned}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_k &= \begin{bmatrix} C_{11} & 0 & 0 \\ 0 & I & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}_k + \\ &\begin{bmatrix} D_1 \\ D_2 \end{bmatrix} u_k + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}_k.\end{aligned}\quad (16b)$$

Here x_2 represents the state variables that give the available y_2 measurements, while the state variables x_1 via C_{11} give the primary variables y_1 that are to be estimated based on the identified input-output relations.

The partitioned innovation form (5) now becomes

$$\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix}_{k+1} = \begin{bmatrix} A_{11} & A_{12} - K_{12} & A_{13} \\ A_{21} & A_{22} - K_{22} & A_{23} \\ A_{31} & A_{32} - K_{32} & A_{33} \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix}_k$$

$$\begin{aligned}
& + \begin{bmatrix} B_1 - K_{12}D_2 \\ B_2 - K_{22}D_2 \\ B_3 - K_{32}D_2 \end{bmatrix} u_k \\
& + \begin{bmatrix} K_{12} \\ K_{22} \\ K_{32} \end{bmatrix} y_{2,k} \quad (17a)
\end{aligned}$$

$$\begin{aligned}
y_{1,k} & = \begin{bmatrix} C_{11} & 0 & 0 \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix}_k \\
& + D_1 u_k + e_{1,k}. \quad (17b)
\end{aligned}$$

We then focus on the specific case where $A_{13} = 0$ and $G_1 = 0$, which means that the process noise v affects the state variables x_1 only through the measured state variables x_2 . This is a realistic assumption in many continuous systems, which will hold approximately also after discretization, provided a high sampling rate.

The asymptotic properties of the Kalman gain are then governed by the following theorem:

Theorem 1 Consider the linear discrete time-invariant model (16) with $A_{13} = 0$ and $G_1 = 0$, or the corresponding continuous model with the same structure. Let v be a stationary or non-stationary multivariable stochastic process, and let $w = \begin{bmatrix} w_1^T & w_2^T \end{bmatrix}^T$ be a stationary multivariable stochastic process with zero mean and diagonal covariance matrices

$$Ew(k+p)w^T(k) = R_w^d \delta(p) = \begin{bmatrix} R_{w_1}^d & 0 \\ 0 & R_{w_2}^d \end{bmatrix} \delta(p) \quad (18)$$

or

$$Ew(t+\tau)w^T(t) = R_w^c \delta(t) = \begin{bmatrix} R_{w_1}^c & 0 \\ 0 & R_{w_2}^c \end{bmatrix} \delta(t), \quad (19)$$

where $\delta(p)$ and $\delta(\tau)$ are the discrete-time and continuous unit impulse functions.

Assume near perfect measurements, with small but possibly different variances μ_i in the measurements. Also assume that all state variables in x_2 are influenced by the process noise v , either directly via G_2 or via the non-measured states x_3 and A_{23} . When all variances $\mu_i \rightarrow 0$, the Kalman gain will then asymptotically include the matrix A_{12} in the following way:

$$K = \begin{bmatrix} K_{11} & A_{12} \\ K_{21} & K_{22} \\ K_{31} & K_{32} \end{bmatrix}. \quad (20)$$

Furthermore, the K_{11} matrix, corresponding to the measurements y_1 , is for a discrete system determined by the solution of the algebraic Riccati equation

$$\begin{aligned}
P_{11} & = A_{11}P_{11}A_{11}^T + A_{12}R_{w_2}^d A_{12}^T \\
& - A_{11}P_{11}C_{11}^T (C_{11}P_{11}C_{11}^T + R_{w_1}^T)^{-1} C_{11}P_{11}A_{11}^T
\end{aligned} \quad (21)$$

and

$$K_{11} = A_{11}P_{11}C_{11}^T (C_{11}P_{11}C_{11}^T + R_{w_1}^T)^{-1}. \quad (22)$$

For a continuous system with the same assumptions concerning A_{13} and G_1 , the K_{11} matrix is determined by the solution of the algebraic Riccati equation

$$\begin{aligned}
0 & = A_{11}P_{11} + P_{11}A_{11}^T + A_{12}R_{w_2}^c A_{12}^T \\
& - P_{11}C_{11}^T (R_{w_1}^c)^{-1} C_{11}P_{11}
\end{aligned} \quad (23)$$

and

$$K_{11} = P_{11}C_{11}^T (R_{w_1}^c)^{-1}. \quad (24)$$

■

For proof of Theorem 1 and some remarks, see [15]. The theorem is also valid for the special case of an OE model. We will then have $K_1 = 0$, $K_{22} = K_{22}^{OE}$ and $K_{32} = K_{32}^{OE}$. The result in Theorem 1 can be combined with the fact that a continuous square system $C(sI - A)^{-1}G$ which is nonsingular in $\text{Re } s \geq 0$ and has covariance matrices $R_v^c = I$ and $R_w^c = \mu I$, approximately has the asymptotic Kalman gain

$$K = \lim_{\mu \rightarrow 0} \left[\mu^{-\frac{1}{2}} G \right] V \quad (25)$$

where V is some orthogonal matrix [9]. The result (25) ignores, however, that also other elements in K than those determined by G will asymptotically approach constant though relatively small values. A combination with Theorem 1 will thus give a more complete solution in the asymptotic perfect measurement case, although restricted to the case where $A_{13} = 0$ and $G_1 = 0$. Two examples of this are given in [15].

A consequence of Theorem 1 is, that given the system (16) with $A_{13} = 0$, $G_1 = 0$ and near perfect noise free y_2 measurements, a reduced model not utilizing all u and y_2 signals may have to be used. Otherwise, numerical problems may occur when the parameter estimates are sought. This is due to decoupling, as illustrated in Example 2. Numerical problems may in fact occur as soon as we have two or more near perfect y_2 measurements. A more important consequence in practice may be that a parsimonious reduced model, even if it is not absolutely necessary from a numerical point of view, may result in less variance in the predictions. In such a case only part of the model will be identified, and as long as there is some noise in one or several of the y_2 signals used as inputs, the parameter estimates will then be biased [10].

3 Experimental dynamic system calibration method

The theoretical analysis in section 2 has shown that numerical identification problems may occur as a result of perfect noise free y_2 measurements. This is not a very likely problem in a practical situation, especially not in an industrial process environment. If it turns out to be a problem, the solution is to leave some of these perfect measurements out, and use a more parsimonious model.

A more important task seen from a practical point of view is to settle for a good set of y_2 measurements to be used as inputs to the identification procedure. This is similar with the problem of finding regressor variables in ordinary least-squares estimation [4]. The inclusion of noisy measurements will give only a limited contribution to the prediction of y_1 , and at the same time the number of unknown parameters to be identified will increase. The total contribution from one or a group of y_2 variables can be explored by use of some added artificial measurement noise or some equivalent method. With a high level of artificial noise on a given y_2 measurement, the influence will be very limited. Gradual reduction of the artificial noise will therefore tend to improve the prediction until the physical noise level is reached.

The final result also depends on choosing a suitable model structure. The optimal choice of input variables and model structure can only be found through proper validation. The following method for identification of the calibration relation (5) is therefore proposed:

1. Perform an informative calibration experiment with only u as input signal and y_1 and y_2 as output signals [11]. Separate the data in one part for identification/calibration and one part for validation.
2. Choose model order and a suitable model structure, possibly by use of a subspace system identification method [12], with u as input and y_1 as output signals. Identify and validate the model, using e.g. the scalar case root mean square error $RMSE = \sqrt{\frac{1}{N} \sum_{k=1}^N (y_{1,k} - \bar{y}_{1,k})^2}$ as a validation criterion.
3. Use one of the y_2 measurements at a time as an input signal together with u , and note the validation improvements for all y_2 signals as expressed by e.g. the $RMSE$ value. Artificial measurement noise reduced in steps may be used for explorative purposes.
4. Include the most informative y_2 measurements as input signals together with u . Choose the number of y_2 measurements to use through validation.
5. Low natural noise levels in some of the y_2 measurements may result in numerical problems causing dramatic changes in the estimation model. If this should occur, remove the last introduced y_2 signal and continue with the next one.
6. Explore the possibilities to use a reduced order model, relying on some or all of the y_2 signals and some of the inputs u , and possibly direct coupling from some other inputs to y_1 . Again, a subspace identification method may be helpful.
7. Repeat (2) to (6) several times, and use the y_2 inputs and model structure that give the best validation results.

Here, it is in order to point to an inherent difficulty in this and similar procedures that uses the same data set for validation and comparison of different models. Due to the randomness of both the modelling and the validation data set, local minima problems and randomized initial parameter values, some models may give better validation results than others in a way that is not generally justified. Extensively repeated use of the same validation set may therefore lead to a model that is specifically adjusted to fit the validation set, which then becomes a part of the total modelling set [13]. We must therefore look for validation differences that can be considered as significant, and extra independent validation sets will certainly be helpful.

4 Simulation results

Simulation studies are undertaken, using `dlsim.m` in the Matlab Control System Toolbox [14], and the prediction error method implemented in `pem.m` in the Matlab System Identification Toolbox [8]. The `pem.m` function identifies the system matrices and the Kalman gain, based on the general innovation model (2), or the partitioned innovation model (5) when the measurements y_2 are also used as input signals. The experimental calibration method in section 3 was tested by simulations based on a continuous system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \end{bmatrix} = \begin{bmatrix} -2 & 1 & 1 & 0 & 0 \\ 2 & -6 & 2 & 2 & 0 \\ 1 & 1 & -3 & 0 & 1 \\ 0 & 2 & 0 & -4 & 0 \\ 0 & 0 & 1 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad (26a)$$

$$\begin{bmatrix} y_1 \\ y_{22} \\ y_{23} \\ y_{24} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \\ x \end{bmatrix} \quad (26b)$$

The system was converted to discrete-time assuming zero order hold elements on the inputs, with a sampling time $T = 0.1$, and discrete process and measurement noise was added. The process noise v_1 and v_2 were independent and normally distributed random signals with zero mean and variance $r_v = 0.1$ directly added to the inputs. The measurements y_1 , y_{22} , y_{23} and y_{24} were assumed to have independent and normally distributed random noise with zero mean and variances $r_1 = 10^{-4}$ and $r_{22} = r_{23} = r_{24} = 0.01$. The prediction model (5) was identified based on simulated time series of length $N = 1000$, using the model (with chosen order n)

$$nn = [0, [n, \dots, n], 0, 0, [n, \dots, n], [1, \dots, 1]] \quad (27)$$

adjusted to the number of inputs used (see [8] for definition of nn). The procedure in section 3 was followed,

except that each identification was repeated in $M = 10$ Monte Carlo runs with independent data sets. In order to avoid local minima, each identification and validation with a given data set was repeated five times with randomized initial parameter values. The calibration experiment in step 1 was performed with the controlled inputs u_1 and u_2 as independent filtered PRBS signal with autocovariance $r_{uu}(p) = 0.5^{|p|}$ (see [6], example 5.11 with $\alpha = 0.5$).

Steps 2 to 4 and step 6 using the marked (x) signals as inputs can be summarized as follows:

step	n	u_1	u_2	y_{22}	y_{23}	y_{24}	$10^4 * RMSE$
2	1	x	x				566 ± 93
2	2	x	x				294 ± 41
2	3	x	x				278 ± 34
2	4	x	x				303 ± 45
2	5	x	x				289 ± 60
3	3	x	x	x			185 ± 18
3	3	x	x		x		196 ± 24
3	3	x	x			x	212 ± 30
4	3	x	x	x			181 ± 36
4	3	x	x	x	x	x	161 ± 31
6	3	x		x	x	x	204 ± 15
6	3		x	x	x	x	139 ± 16
6	2		x	x	x	x	153 ± 23
6	2			x	x	x	212 ± 13

The conclusion from this is that we should use u_2 , y_{22} , y_{23} and y_{24} as inputs and system order $n = 3$ or possibly $n = 2$. As a comparison, the best result with the y_{22} and y_{23} noise levels reduced to $r_{22} = r_{23} = 0.0001$ was found to be $RMSE = 0.0041 \pm 0.0003$, with $n = 1$ and only y_{22} and y_{23} as input signals.

With a short sampling time, the system in this example will be of the type considered in Theorem 1, and numerical problems should therefore be expected with near perfect measurements. However, in order to encounter such problems in step 5 with $N = 1000$, all measurement noise levels had to be decreased to $r = 10^{-15}$ at the same time as the sampling time was reduced to $T = 0.001$.

Identification with only u_1 and u_2 as inputs and y_{22} , y_{23} and y_{24} as outputs was also performed, with the function `canstart.m` in the System Identification Toolbox used for initialization. The results after construction of the predictor (4) were in this case quite discouraging, with frequent failures to find a model, and very inferior validation results when a model was found.

5 Conclusion

The theoretical basis for optimal estimation of nonmeasured primary system outputs y_1 from known inputs u and secondary measurements y_2 is established by use of a partitioned innovation form, with an underlying Kalman filter structure. Identification by use of a rich enough model structure will give an unbiased predictor. Simulations

show that use of y_2 measurements as a basis for estimation of y_1 may give a greatly reduced prediction error. A systematic experimental method is proposed and illustrated by simulation results. The natural next steps in the investigations, will be tests on real industrial data.

References

- [1] Seeborg D.E., Edgar T.F. and Mellichamp D.A., *Process dynamics and control*, Wiley, (1989)
- [2] Mejdell T., Skogestad S., "Composition control of distillation columns using multiple temperature measurements", *AIChE annual meeting in Chicago*, (1990)
- [3] Ljung L., "System Identification", *Technical Report LiTH-ISY-R-1763*, Linköping University, Sweden, (1995)
- [4] Ljung L., *System identification*, Prentice Hall, (1987)
- [5] Lewis F.L., *Optimal estimation*, Wiley, (1986)
- [6] Söderström S., Stoica P., *System identification*, Prentice Hall, (1989)
- [7] Lewis F.L., *Applied optimal control & Estimation*, Prentice Hall, (1992)
- [8] Ljung L., *System identification toolbox for use with Matlab*, The Mathworks Inc., (1991)
- [9] Anderson B.D.O. and Moore J.B., *Optimal control: linear quadratic methods*, Prentice Hall, (1990)
- [10] Söderström T., "Identification of Stochastic Linear Systems in Presence of Input Noise", *Automatica*, **Vol. 17**, 713-725, (1981)
- [11] Goodwin G.C., Payne R.L., *Dynamic system identification*, Academic Press, Inc., (1977)
- [12] Di Ruscio D., "A method for identification of combined deterministic stochastic systems", *The Third European Control Conference (ECC'95)*, (1995)
- [13] Sjöberg J., Ljung L., "Overtraining, regularization and searching for a minimum, with application to neural networks", *International Journal of Control*, **Vol. 62**, 1391-1407, (1995)
- [14] Grace A., Laub A.J., Little J.N., Thompson C.M., *Control system toolbox for use with Matlab*, The Mathworks Inc., (1992)
- [15] Ergon R., Di Ruscio D., "Dynamic system calibration by system identification methods", *Telemark Institute of Technology*, Porsgrunn, Norway, (1997)